

Formative Assessment Tasks as Indicators of Student Engagement for Predicting At-risk Students in Programming Courses

Ashok Kumar VEERASAMY¹, Mikko-Jussi LAAKSO²,
Daryl D’SOUZA³

¹*Department of Software Engineering, LUT University, Lahti, Finland*

²*Department of Computing, University of Turku, Turku, Finland*

³*School of Computer Science, RMIT University, Melbourne, Australia*
e-mail: ashok.veerasamy@lut.fi, milaak@utu.fi, daryl.dsouza@rmit.edu.au

Received: March 2021

Abstract. Previous studies have proposed many indicators to assess the effect of student engagement in learning and academic achievement but have not yet been clearly articulated. In addition, while student engagement tracking systems have been designed, they rely on the log data but not on performance data. This paper presents results of a non-machine learning model developed using ongoing formative assessment scores as indicators of student engagement. Visualisation of the classification tree results is employed as student engagement indicators for instructors to observe and intervene with students. The results of this study showed that ongoing assessment is related to student engagement and subsequent final programming exam performance and possible to identify students at-risk of failing the final exam. Finally, our study identified students impacted by the Covid-19 pandemic. These were students who attended the final programming exam in the semester 2019–2020 and who scored well in formative assessments. Based on these results we present a simple student engagement indicator and its potential application as a student progress monitor for early identification of students at risk.

Keywords: student engagement indicators, student engagement analytics tracking, ongoing assessment tasks, engagement tracking systems.

1. Introduction

Computer programming is a fundamental course in computer science and IT curricula and offered to students from other streams. However, it is widely accepted that computer programming language course is difficult one for novice students (de-la-Fuente-Valentín, Pardo, & Kloos, 2013) (Ali & Smith, 2014) and which affects enrollment and retention for computer science and IT programs (Luxton-Reilly, *et al.*, 2018). Recently, due to COVID-19 pandemic, face-to-face learning moved to e-learning, which in turn

potentially reduced the interaction between teachers and students. As such, there is much discussion about web-based learning. Notably, educators are looking for ways of using data captured and collected via these technologies for measuring student engagement, in order to develop effective instructional interventions to reduce the dropout and failure rates. Student engagement is a key to student achievement and retention (Phan, McNeil, & Robin, 2020). Several studies have been carried out to determine the factors for measuring student engagement (Bond, 2020; Jeongju Lee, 2019). In addition, student progress monitoring systems have been developed to use such indicators for assessing the effect of engagement in student performance (Nagy, 2016; Henrie, pHalverson, & Graham, 2015). However, most of these engagement monitoring systems rely on student demographic and or learning management system *access* data but not on *performance* data. In addition, the factors that serve as student engagement indicators have not yet been clearly articulated and the term engagement is used differently in various contexts. The anecdotal evidence from our teaching staff in computer programming has suggested that, on average only 65% of students enrolled attend the final exam. Furthermore, 77% of students who attended the final exam pass the course. As such, the need to identify students who are not engaged and who need support is crucial for instructors to provide timely interventions before they fully disengaged from studies.

Hence, the objective of this study is to develop a student bi-weekly student engagement indicator model based on student ongoing formative assessment performance for introductory programming courses. Our interest in such a study was prompted by the increasingly prevailing presence of students entering our first-year programming courses from different streams, and the need to develop inclusive teaching strategies to prompt student use of self-regulated learning strategies in programming courses. In addition, this study was motivated by the need for reporting student ongoing assessment progress and performance as indicators of student engagement to support the continuous process of learning. Assessment tasks play vital role in student learning and impacts student performance (Veerasamy, D'Souza, Apiola, Laakso, & Salakoski, 2020). Notably, the results of ongoing formative assessment tasks reveal “a qualitative insight about students learning rather than a score” (Shepard, 2005). In addition, poor or incomplete assessment scores is a sign of disengagement and at risk of dropping out. Furthermore, ongoing assessment results have the potential to identify students who are disengaged in learning (Tempelaar, Rienties, & Giesbers, 2015).

Therefore, this study includes performance in ongoing assessment tasks as key factors for measuring student engagement and to identify at-risk students before final programming exam. Our study uses the following independent variables to measure student engagement in learning programming: homework and demo exercise results during the first eight weeks of the semester. Consequently, we pose and address the following research questions (RQs):

- RQ1.** Is student performance in ongoing assessment tasks related to student engagement and subsequent final programming exam performance?
- RQ2.** Is it possible to propose the student performance in ongoing assessment tasks as a student engagement indicator and as a predictor to identify at-risk students of final programming exam?

Towards addressing these questions, this article is organized as follows. The learning analytics background presents how understanding and examining learning data foster student learning skills. The literature review section presents studies conducted around student engagement and significance of formative assessment tasks in student engagement. The research method section describes the research methodologies used in this study to find the answers for our research questions, including the details about the course. The data analysis and results section present the findings of the study, which we discuss in depth in discussion section. The conclusion and future work (including educational implications) summarises our findings and presents limitations. Finally, the educational implications section presents the contribution of our findings to education including how to develop a more innovative reporting that depicts student engagement over a time period in a variety of visual forms.

1.1. Learning Analytics Background

Learning analytics (LA) is a composition of set of techniques and algorithms used to measure, collect, analyse and to report results on collected data about learners and their contexts, in order to directly support instructors and students (Pardo, 2014). LA is an emerging field that seeks to answer the questions raised around teaching and learning towards enhancing aspects of learning. Educational technologies such as Moodle, ViLLE, and Blackboard to support student learning have offered opportunities to gain insights into teaching and learning via LA (Zhang, Ghandour, & Shestak, 2020; Veerasamy, D'Souza, Apiola, Laakso, & Salakoski, 2020; Hamil, 2020). Higher educational institutions use such analytics data to understand and address student success, retention and graduation rates. There has been research conducted on student learning data to identify patterns of learning behaviors with little human intervention for actionable intelligence to support students within the shortest possible time frame. (Aldowah, Al-Samarraie, & Fauzy, 2019; Liz-Domínguez, Rodríguez, Nistal, & Mikic-Fonte, 2019). However, LA is a moral practice and should be focused on understanding rather than measuring. That is, “how effectively should these results be further mined to provide insights into teaching and learning?” Hence, the objective of this study is to develop a student engagement indicator model to provide detailed information about each student’s ongoing assessment performance for introductory programming courses.

2. Literature Review

2.1. Indicators of Student Engagement

Engagement is the quality of effort students themselves cognitively pursue with their educational activities, which contribute directly to desired learning outcomes (Groccia, 2018). Student engagement is central to academic success in education (Murray, 2018). Engagement is positively linked with academic achievement and student satis-

faction (Tempelaar, Rienties, & Giesbers, 2015; Carolis, D'Errico, Paciello & Palestra, 2019). The reviews of literature on student engagement listed various dimensions of engagement as indicators. They are; behavioural engagement (attend lectures, active participation), emotional engagement (interest, enjoyment, and satisfaction), cognitive engagement (mental effort, goals, and expectations), academic and performance engagement (meets or exceeds assignment requirements, time invested in learning and assessment tasks, assessment task scores and grades). Several studies have examined the impact of engagement on student achievement, attrition, retention, motivation and institutional success (both face-to-face and e-learning) (Curtis R Henrie, 2015; Jeongju Lee, 2019). For example, Kanaparan *et al.* examined the relationship between emotional engagement and student performance in introductory programming courses. They found a positive correlation between enjoyment and student performance, and enjoyment can be considered as indicator of student engagement for introductory programming courses (Kanaparan, Cullen, & Mason, 2019).

Similarly, several studies have explored the value of different data sources as inputs to measure student engagement in multiple courses. These inputs included lecture attendance, number of submissions, number of clicks, number of visits, number of lecture/assessment tasks downloads, number of posts in discussion forums and more. Grey *et al.* conducted a study on measuring lab attendance and student engagement in computer science and reported that attendance has no significant impact on student engagement (Grey & Gordon, 2018). Hussain *et al.* developed a predictive model using various machine learning algorithms with student log data, the highest education level attained, final results, and assessment scores, as inputs to predict low-engagement students, and the relationship between student engagement and course assessment scores. They identified that the number of clicks on a homepage, content and student discussion forums are significantly related to student engagement (Hussain, Zhu, Zhang, & Abidi, 2018). Henrie *et al.* explored the relationship between student activity log data from learning management system (LMS) identified that LMS log data may be used to identify students who need support. However, LMS log data may not be used as a proxy to measure student cognitive and emotional engagement (Henrie, Bodily, Larsen, & Graham, 2018). Azcona *et al.* developed a predictive model to identify at risk students by using the number of submissions and virtual learning environment interaction logs as inputs to predict at-risk students in programming. However, the fail prediction confidence of the model was moderate (64.38%), and the model's prediction results were heavily depended on the last lab exam data held at the end of the semester (Azcona & Smeaton, 2017). Most of these studies used behavioural and cognitive engagement factors only as prime indicators of student engagement, but not on course performance data. In addition, no previous studies have examined or listed a concrete inventory for use as possible indicators for measuring student engagement.

2.2. Student Engagement Tracking Systems

Student engagement tracking systems are used to report how students engage with their studies. The advent of the Internet and educational technologies has spawned

various types of student engagement systems for assessing and displaying student engagement. Moreover, analytics in student engagement has shifted the focus from predicting outcomes to displaying information about student engagement (Verbert, Duval, Klerkx, Govaerts, & Santos1, 2013). This type of tracking systems uses LMS log data, utilisation of course resources, student attendance, assessment scores, behaviour monitoring, and classroom participation, to facilitate visualisation of student engagement (Hussain, Zhu, Zhang, & Abidi, 2018; Burnik, Zaletelj, & Košir, 2017; Basu, Lohani, & Xia, 2019). Most of these tracking systems use machine learning predictive modelling techniques. For example, Burnik *et al.* conducted a study to measure student engagement (attention) in the classroom using facial and body properties of student captured via Kinect sensor-based attention monitoring system as inputs, for predicting (predictive analytics) and visualising student engagement (Burnik, Zaletelj, & Košir, 2017). The review on most common data sources used in dashboard visualisations of student engagement stated that 50% of the dashboards use artefacts produced by learners; time log entries, test scores and self-assessments as indicators of disengagement for predicting at-risk students (Vytasek & Winne, 2020). However, these tracking systems have shortcomings:

- (i) While they present student engagement to relevant stakeholders (student, instructor, course administrator), they do not provide adequate information about how a student is engaging.
- (ii) The sensor reliability of such tracking systems is questionable as the sensor detection depends largely on orientation of human features and may produce erroneous results. Moreover, such sensor-based tracking systems use small subsets of features for training, to measure the prediction accuracy of student engagement.
- (iii) The machine learning based predictive models used in tracking systems do not provide aspects of engagement that are problematic for stakeholders to understand. Most of these systems rely on student presage and LMS log entries but not on student performance data.

2.3. Formative Assessment Data as Indicators of Academic and Performance Engagement

Formative assessment (FA) is conducted by academics during a course, typically aligned with the course syllabus requirements. On the other hand, summative assessment is conducted at the end of the course to evaluate student final course performance. FA results reflect students' incremental progress in learning. FA is aimed at stimulating and directing student learning and plays a significant role in the student learning process (Timmers, Walraven, & Veldkamp, 2015). Continuous FA data has predictive power of identifying disengaged students (Tempelaar, Rienties, & Giesbers, 2015). FA practices enhance student engagement related behaviours (Nichols & Dawson, 2012). Educators use formative assessment tasks such as homework to identify where students are struggling in programming to assist with, and to address their problems (Veeras-

amy, *et al.*, 2016). Several studies have emphasised the importance of continuous FA tasks in the process of learning, engagement and achievement (Holmes, 2018; Hussain, Zhu, Zhang, & Abidi, 2018; Nguyen, Rienties, LisetteToetenel, Ferguson, & Whitelock, 2017). Holmes explored the monitoring of engagement through virtual learning environment to identify the impact of formative and summative assessment tasks on student engagement related behaviours. They reported that the adoption of low-stakes continuous e-assessments into a module increases student engagement (Holmes, 2018). Nguyen *et al.* found that computer-based assessment activities influence student pass rates (Nguyen, Rienties, LisetteToetenel, Ferguson, & Whitelock, 2017). Hussain *et al.* reported instructors can identify low-engaged students from early assessment scores (Hussain, Zhu, Zhang, & Abidi, 2018). These studies revealed that students who underperformed in formative assessment tasks showed signs of disengagement and potentially failed in the final exam.

In summary, as per the aforementioned sub-sections of this paper contributes the following about our study; (i) It attempts to develop a simple non-machine learning based model with explanatory predictor variables, selected based on the previous research findings. (ii) It adopts the simple classification tree analysis method (model) and confusion matrix table based statistical analysis to track students who are not engaged and may fail in the final exam. (iii) It proposes the student performance in ongoing assessment tasks (FAs in this context) as a student engagement indicator for any course with (continuous) formative assessments and a final exam. In addition, to our knowledge no previous study has sought to predict student engagement via a non-machine learning classification tree analysis using a confusion matrix table.

3. Research Methodology

The objective of this study was to identify and visualise students who were not engaged and not qualified to sit, or may fail, the final exam. The homework and demo exercise scores secured during the first two weeks through to eight weeks of the semester were used as source data for measuring student engagement, in order to identify potential indicators of student engagement and to predict students at-risk of failing the final programming exam. Data was collected from 391 students enrolled in the course *Introduction to Programming* over four autumn semesters (years 2016, 2017, 2018 and 2019). The model (Fig. 1) was created based on student performance in continuous FAs (homework and demo) as the context of student engagement. Table 1 presents the dataset collected for this study.

Table 1
Dataset collected for measuring student engagement in programming

Year/semester	2016	2017	2018	2019
Number of students enrolled	93	94	102	102

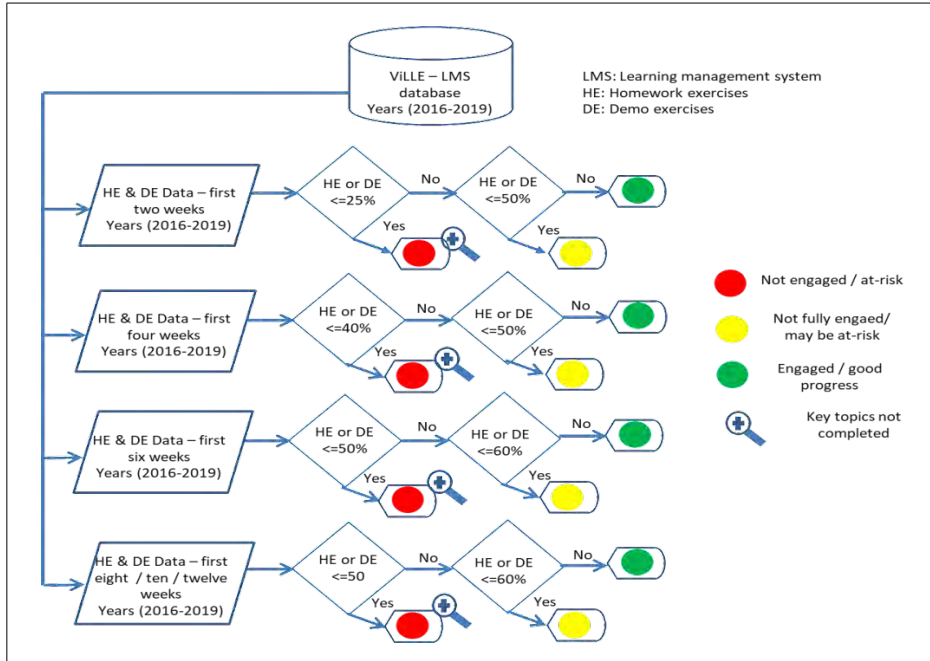


Fig. 1. CTA based model adopted for this study.

The dataset collected for the years 2016–2018 was used to measure the student engagement and to identify at-risk of course failure via a classification tree analysis manually created (CTA), to respond to RQ1. The dataset collected for the year 2019 was then used as test data to answer RQ2. Data was collected via ViLLE and, MS-Excel and R software were used for statistical analysis.

3.1. Overview of the Course and Description of Student Engagement Variables

Introduction to Programming is taught using the Python programming language as a teaching vehicle. It is offered once a year to students from different disciplines. This course is presented in English and the assessment tasks are delivered and graded by the ViLLE-learning management system. The duration of the course is 12 weeks. The assessment tasks include homework, demo exercises, project work and final examination. The final exam is a hurdle and student must secure at least 50% to pass the course. The final course grade is calculated based on scores secured in a final exam and bonus points obtained via formative assessment tasks (homework and demo) and lecture attendance (Veerasamy, 2020).

ViLLE: It is a learning management system / e-learning tool mainly used to deliver, manage lecture notes, formative and summative assessment tasks for programming students (Veerasamy, Daryl D’Souza, & Laakso, 2018).

Homework exercises (HE): These are weekly formative assessment tasks distributed via ViLLE for students to practice and submit their answers electronically over a total of 10–12 weeks. The submitted answers for HE is automatically graded by ViLLE (Veerasamy, Daryl D'Souza, & Laakso, 2018).

Demo exercises (DE): The DE for *Introduction to Programming* was provided weekly for 10 weeks via ViLLE. Students prepare solutions for DE before attending the DE session and a few students are then randomly selected via ViLLE to demonstrate their answers in supervised classes. No marks are awarded for class demonstrations. However, students who complete the DE are instructed to enter their responses in designated (lecturer's) computer to record the number of DEs completed. The marks for DE are calculated by ViLLE, based on their registered responses in the designated computer (Veerasamy, Daryl D'Souza, & Laakso, 2018).

Final exam (FE): This is an online summative assessment conducted using ViLLE at the end of the semester. As noted, FE is mandatory and to be eligible to sit for the FE students must previously have secured at least 50% in HE, 40% in DE, and expected to submit the project work before FE (Veerasamy, Daryl D'Souza, & Laakso, 2018).

The purpose of this study was to notify instructors and students about disengaged students who may at risk of failing the FE, based on their performance in FA tasks in the first eight weeks. So, this study used a simple method; CTA (manually created) that uses FA scores (HE & DE) secured during the first two weeks through to eight weeks of the semester. Fig. 1 illustrates the research methodology adopted for this study to measure and visualise student engagement, and to predict students at-risk of failing the final exam.

As shown in Fig. 1, The HE and DE data for the course was collected after two weeks, four weeks, six weeks and eight-twelve weeks of the years 2016, 2017, 2018 and 2019 used for student engagement indicator model development. The CTA method was created to evaluate and visualise disengaged/at-risk students based on assessment tasks performance in the first two weeks (Week 2), four weeks (Week 4), six weeks (Week 6), and eight-twelve weeks (Week 8) of the semester. In this study we defined students that did not submit in or secured <25% selected formative assessments (HE or DE) in the first two weeks, <=40% in the first four weeks and <=50% in the first six or eight–twelve weeks as disengaged and at-risk students of FE. On the other hand, students that secured >50% in HE or DE in the first two and four weeks, and >60% in the subsequent weeks of the course as engaged and not-at-risk students of FE. The decision tree split values (<=25%, 40%, 50% and 60%) were set based on our computer programming teaching staff suggestions. For example, as noted, student must secure at least 40% in DE and at least 50% in HE to sit for the FE. As such, we defined the decision tree split values to be in between 0%, 25%, 40% and 50% to identify who are not engaged.

In addition, to measure the classification accuracies of the model (Fig. 1) we used confusion matrix table (CMT) analysis to answer RQs. CMT is mainly used to compute prediction sensitivity, specificity in order to weigh the prediction accuracy of the developed models (Mueen. Zafar & Manzoor, 2016; Liao *et al.*, 2019). Table 2 shows the skeleton of binary classification problem-based confusion matrix table used for this study.

Table 2
Confusion matrix table for performance measurement of CTA

		Predicted by CTA	
		Red + Yellow (Not engaged)	Green (Engaged)
Actual	Not engaged	True positive	False positive
	Engaged	False negative	True negative

True positive (TP): In this study, the TP value represents the number of disengaged students (Red + Yellow) who are correctly identified by the CTA as at-risk students of FE failure.

False positive (FP): In this study, the FP value represents the number of engaged students (Green) who are not correctly identified by the CTA as not-at-risk students of FE failure.

True negative (TN): In this study, the TN value represents the number of engaged students (Green) who are correctly identified by the CTA as not-at-risk students of FE failure.

False negative (FN): In this study, the FN value represents the number of disengaged students (Red + Yellow) who are not correctly identified by the CTA as at-risk students of FE failure.

The disengaged and at-risk students of FE prediction accuracy (sensitivity) based on CTA results was measured by:

$$Sensitivity = \frac{TP}{(TP + FN)}$$

The engaged and not-at-risk students of FE failure prediction accuracy (specificity) based on CTA results was measured by:

$$Specificity = \frac{TN}{(TN + FP)}$$

The overall student engagement prediction accuracy was calculated as the number of correct predictions made by the CTA by the total number of actual values and multiplied by 100 to get the prediction accuracy.

4. Data Analysis and Results

The methodology (Fig. 1.) defined for this study posits that students who underperform in the ongoing FAs (HE and or DE) will be considered as disengaged students and may fail in FE. So, first we analysed the data collected for this study. Table 3 presents the initial results of data analysis.

Table 3
Number of students failed the course in the years 2016–2019

Year	Number of students enrolled	Number of students failed the course [did not attend + secured <50 marks in FE]	Number of students Failed in FE (in %)
2016	93	19 + 27 = 46	49.46
2017	94	26 + 11 = 37	39.36
2018	102	35 + 23 = 58	56.86
2019	102	50 + 6 = 56	54.90
Total	391	197	50.38

As noted, in total, 391 students enrolled in 2016–2019 and of these 197 students received fail grade in introductory programming course. Notably, 130 students did not attend the FE in the period 2016–2019. The preliminary findings indicated that nearly 40%–50% students failed the course every year. In addition, 102 students enrolled in the year 2019 and of these nearly 50% of students did not attend the FE in the year 2019 (Table 3) and this was addressed in discussion section.

As noted, one of the objectives of this study was to measure student engagement based on student performance in ongoing FAs during the semester. So, we used CTA on 2016–2018 data to answer RQ1. Table 4 presents the total number of disengaged, not fully engaged, and engaged students based on their performance in ongoing continuous assessment tasks in Week 2–Week 8 of the period 2016–2018.

In total, 289 students enrolled in an introductory programming course for the years 2016–2018. Of these, 141 students did not complete the course with a passing grade. Our CTA results show that on average 84% of students who identified as disengaged (Red) in the first eight weeks received a fail grade in the course. Similarly, 54% of students that partially engaged (Yellow) in assessments task completion during the first eight weeks did not receive pass grade in the course. On the other hand, on average, nearly

Table 4
Student engagement in formative assessment tasks results (Week 2–Week 8)
of the period 2016–2018

2016–2018	Red			Yellow			Green		
	Not engaged	FE Fail	Prediction accuracy (%)	Not fully engaged	FE Fail	Prediction accuracy (%)	Engaged	FE Fail	Prediction accuracy (%)
Week 2	64	52	81.25	81	43	53.09	144	46	68.06
Week 4	76	63	82.89	15	9	60.00	198	69	65.15
Week 6	95	79	83.16	28	16	57.14	166	46	72.29
Week 8	104	91	87.50	28	13	46.43	157	37	76.43
Average prediction accuracy:	83.70%			54.16 %			70.48 %		

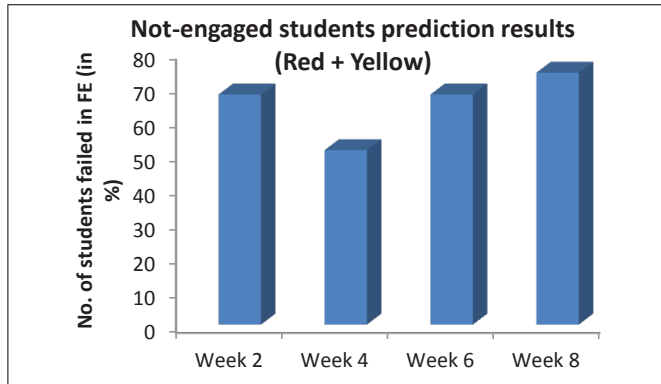


Fig. 2. CTA results: 2016–2018.

71% of students who identified as engaged in formative assessment tasks (Green) pass the course (Table 4).

To answer RQ1, we aggregated the results of disengaged and partially disengaged (Red + Yellow) to identify the relation between student performance in FAs, engagement and final programming exam results. Fig. 2 presents the results of disengaged students (Red + Yellow) that failed in FE in the period 2016–2018.

Nearly 67% of students that not fully engaged in the first two weeks assessment task completion failed in FE. Similarly, on average, 70% of students that were identified as disengaged in Week 6–Week 8 did not pass the FE. However, only 51% of students that not engaged in the first four weeks of assessment task completion failed in FE. This should be further analysed.

To confirm the validity of the aforementioned results we used confusion matrix table (CMT) based statistical analysis. The prediction results (Table 4 and Fig. 2) were evaluated based on CMT computed via R coding to answer RQ1 and to proceed for testing it on 2019 data. That is, student performance in ongoing FA tasks is related to student engagement and subsequent final programming exam performance. Table 5 shows the CMT results of student engagement for the period 2016–2018.

The CMT results for Week 2 data revealed that 67% (sensitivity) of students who secured low scores ($\leq 50\%$) in FAs failed in FE. However, the CMT results of CTA based

Table 5
CMT results on CTA based model for the period 2016–2018

2016–2018	Overall prediction accuracy	Sensitivity (Red + Yellow) (disengaged)	Specificity (Green) (engaged)	95% of CI
Week 2	66.78	67.38	66.22	61.03–72.19
Week 4	69.55	51.06	87.16	63.89–74.80
Week 6	74.39	67.38	73.83	68.95–79.33
Week 8	77.51	73.76	81.08	72.75–82.19

on first four weeks of assessment task performance data shows that, only 51% of students who secured $\leq 50\%$ in FAs failed in FE. It should be further analysed. On the other hand, on average, nearly 71% of students that secured low scores in Week 6–Week 8 failed in FE. Similarly, on average, 77% of students that secured high scores during Week 2–Week 8 in the FAs identified as engaged pass the FE. These results imply that there is a relation between student performance in ongoing assessment tasks, student engagement and on subsequent FE performance. Hence, we tested the model (Fig. 1) on 2019 data for generalisation. Fig. 3 presents the CTA results for the year 2019.

The CTA on 2019 data yielded mixed results. On average, 79% of students that were identified as disengaged (Red) based on FA performance in the first on first two / four / six / eight weeks of the semester 2019 failed in FE. However, on average only 30% of students that were identified as partially engaged (Yellow) in Week 1–Week 8 assessment task completion failed in FE. Similarly, 49% of students (on average) those identified as engaged (Green) failed in FE despite secured good scores in FAs (Fig. 3). We also applied CMT statistical analysis on CTA results of 2019 to confirm the validity of the CTA results. Table 6 presents the CMT results for the year 2019.

Our CMT based statistical results revealed that the overall student engagement prediction accuracy on identifying student engagement is 56%. Similarly, the average of disengaged and at-risk student of FE prediction accuracy (sensitivity) is poor (35%). On the other hand, the average of engaged and not-at-risk student of FE failure prediction accuracy is high but with many false negatives (82%). As the CTA and CMT statistical results on identifying disengaged and engaged students on 2019 data were not significant, we reanalysed 2019 FE results against FAs performance. We found that in total, 56 students did not pass FE in 2019. Of these, 50 students did not sit the

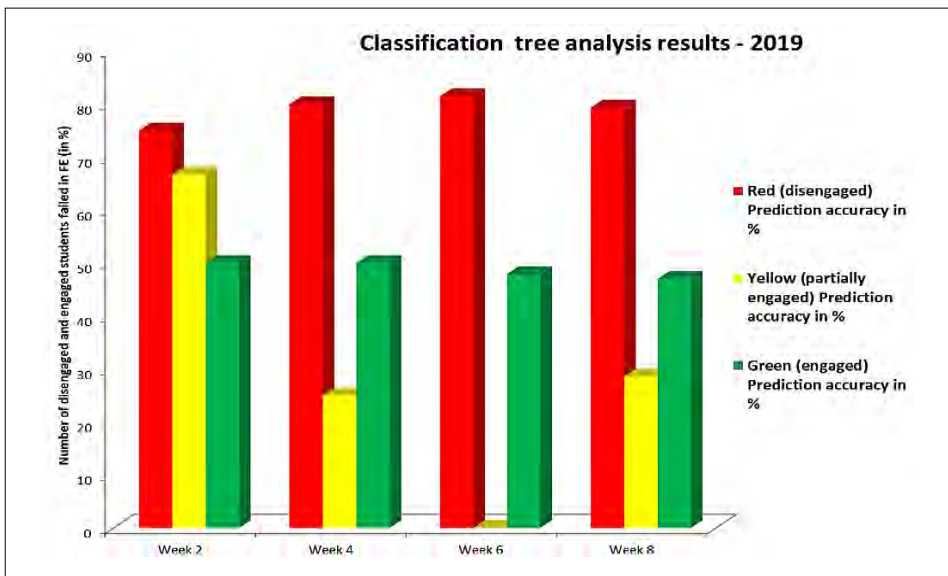


Fig. 3. CTA results of student engagement for the year 2019.

Table 6
CMT results on CTA based model for the year 2019

2019	Overall prediction accuracy	Sensitivity (Red + Yellow) (Disengaged)	Specificity (Green) (Engaged)	95% of CI
Week 2	54.9	28.57	86.96	44.74–64.78
Week 4	54.9	30.36	84.78	44.74–64.78
Week 6	54.9	37.29	79.07	44.74–64.78
Week 8	58.82	44.64	76.09	48.64–68.48

FE. Moreover, nearly 27 students (out of 50 students who did not attend the FE) that secured high scores in FAs during the semester and identified as engaged (Green) (from Week 2 through Week 8) did not attend the FE. So, we initiated a 45-minute discussion meeting with a lecturer of the course and the outcomes of meeting were addressed in discussion section.

5. Discussion

The main objective of this study was to develop a bi-weekly student engagement indicator based on ongoing assessment tasks performance and subsequently to identify students that need support for better teaching interventions. We analysed the data set of the semesters 2016–2019 extracted from ViLLE for selected programming course. We used formative assessment performance data to measure student engagement in learning programming and to identify student at-risk of FE. To answer our RQs we used the statistical CTA (Fig. 1) and CMT (Table 2).

The CTA results of 2016–2018 revealed that students who underperform in assessment tasks or low engagement in assessment task completion (Red) may skip or fail in the FE (84%). Similarly, students that complete assessment tasks or high engagement (Green) in assessment task completion may pass the FE (71%). These results suggest that formative assessment scores leave signs of disengagement (Table 4).

The CMT results on measuring student engagement and identification of at-risk students yielded good results and answered RQ1. The overall prediction accuracy results of Week 2–Week 8 (2106–2018) suggest that it is possible to measure student engagement based on ongoing FAs results (Table 5). The confusion matrix results for disengaged and at-risk students of FE prediction (sensitivity) produced good results (67%: Week 2 and 74%: Week 8) with a small downward spike (51%: Week 4) over the period of 12 weeks. Similarly, the results of specificity also had gradual increase in identifying engaged and not-at-risk students of FE failure (66%: Week2 and 81%: Week 8) with a small downward spike (74%: Week 6). The decline in Week 4 (sensitivity: 51%) and Week 6 (specificity: 74%) student engagement results should be further analysed. Hence, these results answered our RQ1. The CTA and confusion matrix based statistical results imply that learning is dynamic so, there is a need to measure student engagement throughout the course period to track disengaged students in order to provide academic support before

they fall. Furthermore, these findings support the conclusion arrived at by Tempelaar *et al.*, Hussain *et al.*, and Carolis *et al.* (Tempelaar, Rienties, & Giesbers, 2015; Hussain, Zhu, Zhang, & Abidi, 2018; Carolis, D'Errico, Paciello, & Palestra, 2019).

As noted, the CTA and CMT results of 2019 produced mixed results. The statistical results of CTA based model (Fig. 1) of the year 2019 showed that it is possible to predict disengaged and at-risk students of FE based on student performance in selected FAs (Fig. 3). In particular, the results of disengaged (Red) during Week 2–Week 8 implies that students who did not complete or secure low scores in FAs would fail in FE (79% average). Moreover, these results support the presence of an association between student performance in ongoing assessment tasks, student engagement, and subsequent performance in the final exam (RQ1). However, the results relating to identifying partially disengaged (30% average) and engaged (Green: 49% average) produced insignificant results. Similarly, the CMT results of 2019 produced insignificant results (Table 6). On average, the overall prediction accuracy on measuring student engagement was 56% only. In addition, the results of sensitivity (35% on average) did not support the statement “Student performance in FAs leaves signs of disengagement in learning”. On the other hand, the engaged and not-at-risk students of FE failure prediction accuracy (specificity) yielded significant results (82% on average) in identifying engaged students. However, nearly 49% of students who completed and secured high scores in FAs throughout the semester did not attend the FE. Hence, we reanalysed the FE results of 2019 and had a meeting with the lecturer who taught that year. The outcomes were:

- (i) The FE for the autumn semester 2019 was conducted four times (two times in December 2019, end of January 2020, and in the first week of March 2020). As noted earlier, in total 52 students only attended the FE in 2019. Of these, 46 students attended the FE in December 2019.
- (ii) The course had many exchange students (numbers not known at time of writing).
- (iii) Due to Christmas and New Year 2020 many students opted to take the FE after January 2020 but failed to register due to the Covid-19 issue.
- (iv) There were no changes to lecture content, but more exercises were added for DE.

These findings prompted speculation that Covid-19 pandemic might have prevented students from attending the FE despite attaining good scores in selected FAs and identified as engaged (Green); this should be further analysed. Despite these mixed results, the results of identifying disengaged and at-risk students (Red) of 2016–2019 implies that student performance in formative assessments may be used as indicators to measure student engagement. Furthermore, the CMT results of 2016–2018 endorse the importance of formative assessments in identifying disengaged and at-risk students of FE.

In addition, from these results the following points emerged. First, formative assessments are designed to evaluate student learning outcomes. However, it is important that instructors take responsibility for utilising analytics tools to shape student outcomes, by monitoring and evaluating student learning outcomes. Furthermore, it assists them to identify which students need support to focus on relevant activities. Second, the results for the period 2016–2018 and the year 2019 provide an answer to RQ2 which is reiter-

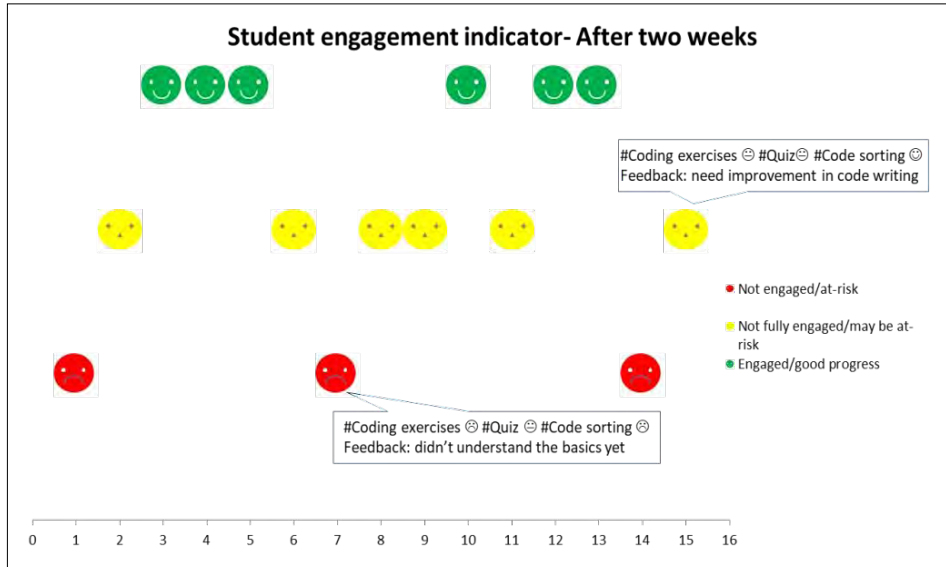


Fig. 4. Student engagement indicator based on first two weeks FA performance.

ated here: “Is it possible to propose the student performance in ongoing assessment tasks as a student engagement indicator and as a predictor to identify at-risk students of final programming exam?” As we concluded earlier, student engagement in continuous assessment tasks does influence student engagement in learning and academic performance (Table 4 and Table 5). Based on the past research findings (Veerasamy, 2020) and results of our study we present a sample screenshot of the student engagement/student progress indicator based on FA performance after two weeks. Fig. 4 illustrates the results of CTA based model (n = 15 in random as an example).

Finally, “How might a CTA based model (Fig. 1) adopted for this study deployed for other courses, if the goal of the instructor is to identify students who are not engaged and or at-risk of course failure?” The research methodology and results of this study may be used by instructors to measure student engagement in other courses. For example, the results of CTA based model may be used in other programming and non-programming courses that use ongoing formative and continuous summative assessment tasks and a final exam, if the goal of the instructor is to measure student engagement and to predict students at-risk of course failure.

6. Conclusions and Educational Implications

Our study involving the use of a CTA and a confusion matrix table showed that results of ongoing assessment tasks (2016–2018) present symptoms of disengagement in studies and may be used as student engagement indicators to identify students who need support. The course lecturer’s report on disengaged and at-risk students of FE prediction accuracy

results of 2019 gave rise to speculation that the Covid-19 pandemic may have impacted students' decisions on attending final programming exam. This needs further analysis.

Our study has some limitations. First, the results were derived from simple a CTA and applied to a specific university course. Second, the results of 2019 did not sufficiently support to a generalization of our findings. Third, we used the first eight weeks of assessment results but, as noted earlier, learning is dynamic. Hence, the results of CTA should be visualized from Week 1 onward and until the end of the course lecture period; this would enable early interventions before at-risk students become disengaged from their studies. This process should be continued every two weeks till the end of the semester. Fourth, this study did not use predictive modelling based on machine learning algorithms, a consideration to be pursued in a future study. Finally, this study did not use the other student engagement factors such as, login data, lecture attendance, and boredom to measure student engagement.

Despite these limitations, the results of this study provide evidence that student performance in ongoing assessments may influence student engagement and subsequent summative assessment performance. Therefore, we intend to seek answers to the following research questions: "What kind of report visualization would help instructors to understand the level of student engagement in learning to focus on related supporting activities?" "How might instructors use the results of student engagement to shape student learning outcomes?" Educators face constant challenges in utilizing such tracking tools for providing timely feedback, amidst the increasing demands and expectations of students. Finally, some instructors were not able to interpret the meaning of graphical representation of data provided via these tools. Hence, we plan to conduct further research in visualization and the challenges faced by educators on using student progress monitoring tools.

The aforementioned limitations aside, our findings contribute to education in the following ways. First, it is possible to measure student engagement in learning and categorize on the basis of continuous assessment task performance. Second, identifying student engagement levels (disengaged, not fully engaged, engaged) may assist instructors to track students' academic progress for decision-making and to predict students who are academically at-risk stage. Third, student engagement and learning analytics go hand in hand together. As such, utilizing learning analytics for tracking student engagement may be useful to find its relationship with learning and academic achievement. For example, in this study we proposed a simple student engagement indicator (Fig. 4) for instructors to track student engagement in ongoing assessment tasks to enhance active learning. Fourth, we used "measurement and reporting of performance data about learners" to evaluate student engagement, which is a subset of learning analytics. It can be further extended for prescriptive analytics to connote something more than traditional performance statistics. Finally, today's learning environments have radically shifted from traditional to interactive, and more online oriented. So, developing student progress tracking systems with feedback may assist instructors to implement learning analytics tools to get actionable data for analysis in order to serve meaningful results to students. However, researchers suggest data served by instructors to students should not discourage them from learning. Therefore, instructors should take responsibility to utilize student engagement indicators/tracking systems properly to shape student outcomes.

References

- Aldowah, H., Al-Samarraie, H., & Fauzy, W.M. (2019, April). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, 37, 13–49. Retrieved from <https://doi.org/10.1016/j.tele.2019.01.007>
- Ali, A., & Smith, D. (2014). Teaching an Introductory Programming Language. *Journal of Information Technology Education: Innovations in Practice*, 13, 57–67.
- Azcona, D., & Smeaton, A.F. (2017). Targeting At-risk Students Using Engagement and Effort Predictors in an Introductory Computer Programming Course. In: D.H.-S. Lavoué É. (Ed.), *European Conference on Technology Enhanced Learning*. 10474, pp. 361–366. Springer, Chham. https://doi.org/10.1007/978-3-319-66610-5_27
- Basu, D., Lohani, V.K., & Xia, K. (2019). Analysis of Students' Personalized Learning and Engagement within a Cyberlearning System. *126th Annual conference and exposition*. American Society for Engineering Education.
- Bond, M. (2020). Facilitating student engagement through the flipped learning approach in K-12: A systematic review. *Computers & Education*, 151, 1–30. <https://doi.org/10.1016/j.compedu.2020.103819>
- Burnik, U., Zaletej, J., & Košir, A. (2017). Kinect based system for student engagement monitoring. *2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)* (pp. 1229–1232). Kiev: IEEE. DOI:10.1109/UKRCON.2017.8100449.
- Carolis, B.D., D'Errico, F., Paciello, M., & Palestra, G. (2019). Cognitive Emotions Recognition in e-Learning: Exploring the Role of Age Differences and Personality Traits. *International Conference in Methodologies and intelligent Systems for Technology Enhanced Learning* (pp. 97–104). Springer. https://doi.org/10.1007/978-3-030-23990-9_12
- Curtis R Henrie, R.B. (2015). Exploring intensive longitudinal measures of student engagement in blended learning. *International Review of Research in Open and Distributed Learning*, 16(3), 131–155. <https://doi.org/10.19173/irrodl.v16i3.2015>
- de-la-Fuente-Valentín, L., Pardo, A., & Kloos, C.D. (2013). Addressing drop-out and sustained effort issues with large practical groups using an automated delivery and assessment system. *Computers & Education*, 61(February), 33–42.
- Grey, S., & Gordon, N. (2018). Approaches to Measuring Attendance and Engagement. *New Directions in the Teaching of Physical Sciences*, 13(1). <https://doi.org/10.29311/ndtps.v0i13.2767>
- Groccia, J.E. (2018). What Is Student Engagement? In *New Directions for Teaching and Learning* (pp. 11–20). Wiley Periodicals, Inc. DOI: 10.1002/tl.20287
- Hamil, D. (2020, October 31). Using Blackboard (VLE) to Support Teaching Practice of Academic Staff in Response to COVID-19. *All Ireland Journal of Higher Education*, 12(3), 1–9. Retrieved from <https://www.tcd.ie/CAPSL/>
- Henrie, C.R., Bodily, R., Larsen, R., & Graham, C.R. (2018). Exploring the potential of LMS log data as a proxy measure of student engagement. *Journal of Computing in Higher Education volume*, 30, 344–362. <https://doi.org/10.1007/s12528-017-9161-1>
- Henrie, C.R., R. Halverson, L., & Graham, C.R. (2015, December). Measuring student engagement in technology-mediated learning: A review. *Computers & Education*, 90(1), 36–53. DOI: 10.1016/j.compedu.2015.09.005
- Holmes, N. (2018). Engaging with assessment: Increasing student engagement through continuous assessment. *Active Learning in Higher Education*, 19(1), 23–34. <https://doi.org/10.1177/1469787417723230>
- Hussain, M., Zhu, W., Zhang, W., & Abidi, S.M. (2018). Student Engagement Predictions in an e-Learning System and Their Impact on Student Course Assessment Scores. *Computational Intelligence and Neuroscience*, 2018, 1–21. <https://doi.org/10.1155/2018/6347186>
- Jeongju Lee, H.-D. S. (2019). Exploring factors, and indicators for measuring students' sustainable engagement in e-learning. *Sustainability*, 11(4), 1–12. DOI: 10.3390/su11040985
- Kanaparan, G., Cullen, R., & Mason, D. (2019). Effect of Self-efficacy and Emotional Engagement on Introductory Programming Students. *Australasian Journal of Information Systems*, 23, 1–23. <https://doi.org/10.3127/ajis.v23i0.1825>
- Liao, S.N., Zingaro, D., Thai, K., Alvarado, C., Griswold, W.G., & Porter, L. (2019, June). A Robust Machine Learning Technique to Predict Low-performing Students. *ACM Transactions on Computing Education (TOCE)*, 19(3), 1–19. DOI: 10.1145/3277569

- Liz-Dominguez, M., Rodríguez, M.C., Nistal, M.L., & Mikic-Fonte, F.A. (2019). Predictors and early warning systems in higher education-A systematic literature review. *LASI-SPAIN*, (pp. 84–99).
- Luxton-Reilly, A., Simon, Albluwi, I., Becker, B.A., Giannakos, M., Kumar, A.N., . . . Szabo, C. (2018). Introductory Programming: A Systematic Literature Review. *Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education* (pp. 55–106). Laranca: ACM. DOI: 10.1145/3293881.3295779
- Mueen, A., Zafar, B., & Manzoor, U. (2016). Modeling and Predicting Students' Academic Performance Using Data Mining Techniques. *I.J. Modern Education and Computer Science*, 11, 36–42. DOI: 10.5815/ijmecs.2016.11.05
- Murray, J. (2018, September 3). Student-led action for sustainability in higher education: a literature review. *International Journal of Sustainability in Higher Education*, 1095–1110. DOI: 10.1108/IJSHE-09-2017-0164
- Nagy, R. (2016). Tracking and Visualizing Student Effort: Evolution of a Practical Analytics Tool for Staff and Student Engagement. *Journal of Learning Analytics*, 3(2), 165–193. DOI: 10.18608/jla.2016.32.8
- Nguyen, Q., Rienties, B., LisetteToetenel, Ferguson, R., & Whitelock, D. (2017). Examining the designs of computer-based assessment and its impact on student engagement, satisfaction, and pass rates. *Computers in Human Behavior*, 76, 703–714. <https://doi.org/10.1016/j.chb.2017.03.028>
- Nichols, S.L., & Dawson, H.S. (2012). Assessment as a Context for Student Engagement. In *Research on Student Engagement* (pp. 457–477). Springer. https://doi.org/10.1007/978-1-4614-2018-7_22
- Pardo, A. (2014). Designing Learning Analytics Experiences. In J. Larusson, & B. White (Eds.), *Learning Analytics From Research to Practice* (pp. 15–35). New York: Springer.
- Phan, T., McNeil, S.G., & Robin, B.R. (2020). Students' patterns of engagement and course performance in a Massive Open Online Course. *Computers & Education*, 95, 36–44. <https://doi.org/10.1016/j.compedu.2015.11.015>
- Shepard, L.A. (2005). Formative assessment: Caveat emptor. *ETS Invitational Conference The Future of Assessment: Shaping Teaching and Learning*. New York, pp. 279–303.
- Tempelaar, D.T., Rienties, B., & Giesbers, B. (2015). In search for the most informative data for feedback generation: Learning analytics in a data-rich context. *Computers in Human Behavior*, 47, 157–167. <https://doi.org/10.1016/j.chb.2014.05.038>
- Timmers, C.F., Walraven, A., & P.Veldkamp, B. (2015, March 9). The effect of regulation feedback in a computer-based formative assessment on information problem solving. *Computers & Education*, 87(9), 1–9. <https://doi.org/10.1016/j.compedu.2015.03.012>
- Veerasamy, A.K. (2020). *Predictive Models as Early Warning Systems For Student Academic Performance in Introductory Programming*. University of Turku, Department of Computing, Turku: TUCS. Retrieved February 18, 2021, from <https://www.utupub.fi/bitstream/handle/10024/150782/TUCSD259.digi.pdf?sequence=1&isAllowed=y>
- Veerasamy, A.K., Daryl D'Souza, R.L., & Laakso, M.-J. (2018). The impact of prior programming knowledge on lecture attendance and final exam. *Journal of Educational Computing Research*, 56(2), 226–253. <https://doi.org/10.1177/0735633117707695>
- Veerasamy, A.K., D'Souza, D., Apiola, M.-V., Laakso, M.-J., & Salakoski, T. (2020). Using early assessment performance as early warning signs to identify at-risk students in programming courses. *2020 IEEE Frontiers in Education Conference (FIE)* (pp. 1–9). Uppsala, Sweden: IEEE. DOI: 10.1109/FIE44824.2020.9274277
- Veerasamy, A.K., D'Souza, D., Lindén, R., Kaila, E., Laakso, M.-J., & Salakoski, T. (2016). The Impact of Lecture Attendance on Exams for Novice Programming Students. *International Journal of Modern Education and Computer Science (IJMECS)*, 8(5), 1–11. DOI: 10.5815/ijmecs.2016.05.01
- Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos1, J.L. (2013). Learning Analytics Dashboard Applications. *American Behavioral Scientist*, 57(10), 1500–1509. DOI: 10.1177/0002764213479363
- Vytasek J.M., A. P., Winne, P.H. (2020). Analytics for Student Engagement. In: M. Virvou, E. Alepis, G.A. Tshirintzis, & L.C. Jain, *Machine Learning Paradigms: Advances in Learning Analytics* (pp. 23–48). Springer. <https://doi.org/10.1007/978-3-030-13743-4>
- Zhang, Y., Ghandour, A., & Shestak, V. (2020, October). Using Learning Analytics to Predict Students Performance in Moodle LMS. *International Journal of Emerging Technology in Learning*, 15(20), 102–115. Retrieved from <https://www.learntechlib.org/p/218306/>

A.K. Veerasamy is a University lecturer cum Researcher at Lapeenranta University of Technology, Lathi, Finland. He did his PhD in Computer Science at University of Turku, Finland. He did his degree of Graduate Certificate in Tertiary Teaching and Learning at RMIT University, Australia. He is an accomplished CS / IT Educator with over 15 years of international experience at universities, colleges, schools and private institutions teaching students from various social and cultural backgrounds. He has also published several articles related to computer science education and learning analytics at ranked conferences and journals. His research and publication interests lie in the area of computer science & IT education, educational technologies (ViLLE) in teaching and learning, computational aspects of learning analytics and computational thinking for STEM education.

M.-J. Laakso is an Associate Professor at Department of Computing, University of Turku, Finland. He is also the director of the Center for Learning Analytics at University of Turku, Finland. His main research interests are Learning analytics, Computer assisted learning, Math & Programming education, Gamification, Learning design, Machine learning & AI in education. He has 20 years of experience in university education research-based development through educational technology solutions. He has published more than 100 international peer-reviewed articles and collected more than 4M€ in R&D projects. The center is developing and awarded by UNESCO as #1 digital learning platform in Finland: ViLLE – the collaborative education tool.

D. D'Souza is an RMIT University Associate, an honorary position. He was an academic member of the School of Computer Science & IT, RMIT University, Melbourne, since 1983, and contributed substantially to all facets of university life, including teaching, research, leadership and community engagement. He holds a PhD (2005) in Computer Science (information retrieval). He has extensive teaching experience at both undergraduate and postgraduate coursework levels and was largely involved in first year teaching of programming. This latter activity became an important part of his reflective teaching practice and scholarly activity in the area of computing education research and since 2011, has had a research interest in computing education research, which includes learning analytics. He has published in computing education research since 2011, has chaired national computing education conferences, and has been an investigator in several research grants to investigate curriculum and pedagogy issues around novice programming.